

예제 2

- 자료: 우리나라에 입국한 관광객 수
- 분석모형
 - 계절형 ARIMA 모형
 - 회귀모형에 의한 추세계절모형 + ARMA 오차 모형
- Training data(tourist.txt): 1981.1~1991.12
- Test data(tour92.txt): 1992.1~1992.12
 - 두 모형의 예측 정확성 측도 비교
- 자료 준비

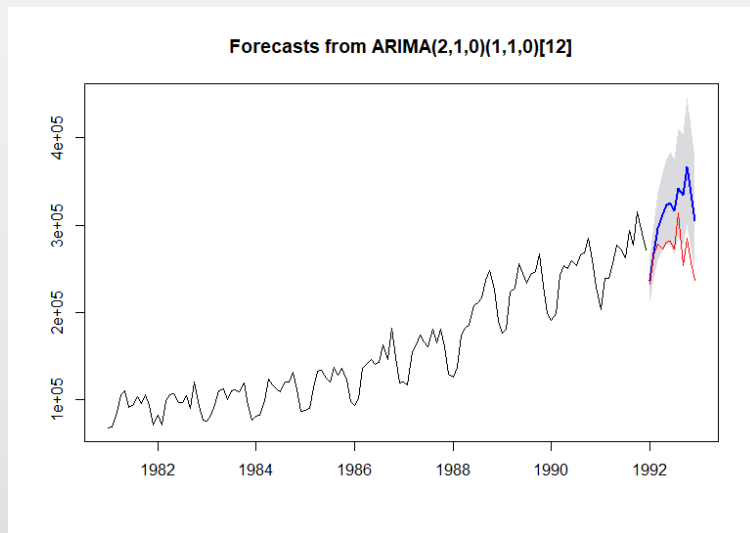
```
> tour <- scan("D:/Data/tourist.txt")  
> tour.ts <- ts(tour, start=1981, frequency=12)
```

```
> tour92 <- scan("D:/Data/tour92.txt")  
> tour92 <- ts(tour92, start=1992, freq=12)
```

```
> lntour <- log(tour.ts)
```

1. 계절형 ARIMA 최종 모형

```
> fit4_1 <- Arima(tour.ts,order=c(2,1,0),  
                  seasonal=list(order=c(1,1,0),period=12),lambda=0)  
  
> fore_arima <- forecast(fit4_1,h=12,level=95)  
> plot(fore_arima)  
> new_t <- seq(1992,by=1/12,length=12)  
> lines(new_t,tour92,col="red")
```



2. 계절추세회귀모형 + ARMA 오차 회귀모형

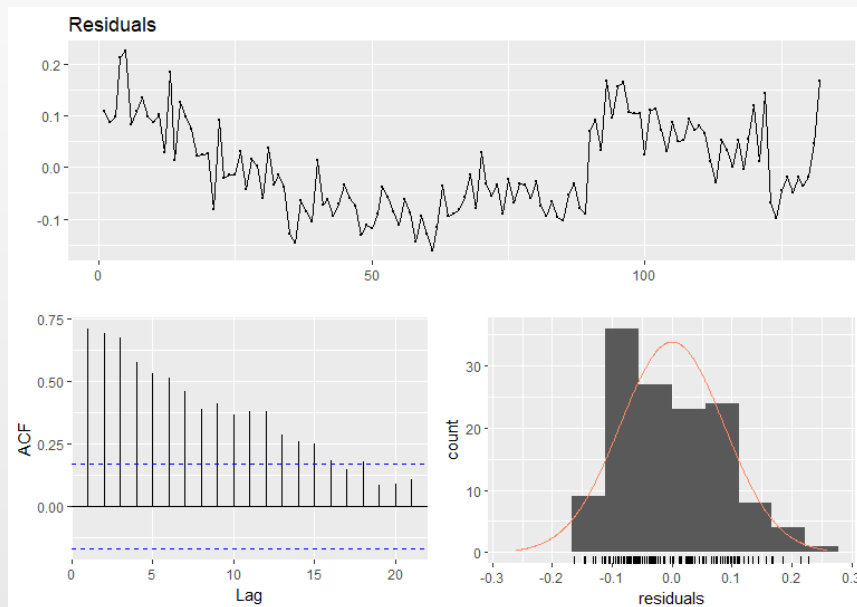
- 1차 추세계절 회귀모형 적합

```
> Time <- time(lntour)
> Month <- cycle(lntour)
> fit1 <- lm(lntour~Time+factor(Month)+0)
```

```
> summary(fit1)
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
Time              0.1195    0.0025   47.78  <2e-16 ***
factor(Month)1   -225.6597    4.9657  -45.44  <2e-16 ***
factor(Month)2   -225.6373    4.9659  -45.44  <2e-16 ***
factor(Month)3   -225.4362    4.9661  -45.40  <2e-16 ***
factor(Month)4   -225.3477    4.9663  -45.38  <2e-16 ***
factor(Month)5   -225.3286    4.9665  -45.37  <2e-16 ***
factor(Month)6   -225.3848    4.9667  -45.38  <2e-16 ***
factor(Month)7   -225.3999    4.9669  -45.38  <2e-16 ***
factor(Month)8   -225.3277    4.9671  -45.36  <2e-16 ***
factor(Month)9   -225.3804    4.9673  -45.37  <2e-16 ***
factor(Month)10  -225.2771    4.9675  -45.35  <2e-16 ***
factor(Month)11  -225.4349    4.9677  -45.38  <2e-16 ***
factor(Month)12  -225.6338    4.9679  -45.42  <2e-16 ***
```

- 잔차 분석

```
> checkresiduals(fit1)
```



- 2차 추세 모형?
- 강한 상관관계?

- 2차 추세모형 적합 시도

- 2차 추세 적합

```
> fit2 <- lm(lntour~Time+I(Time^2)+factor(Month)+0)
```

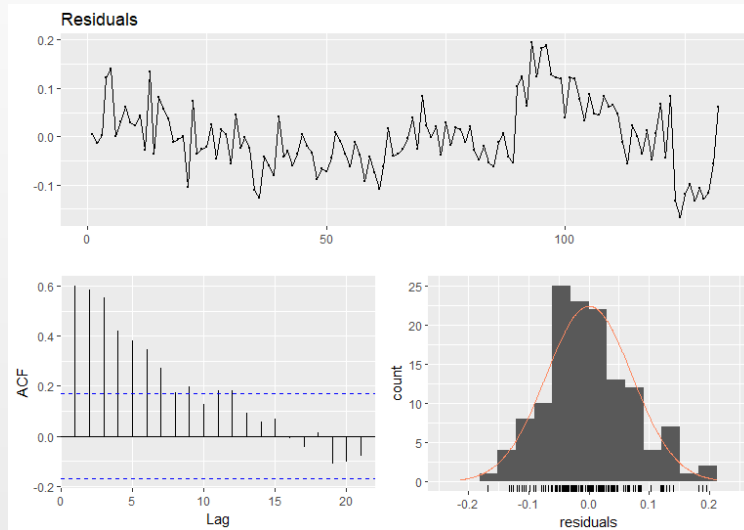
Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
Time	-2.143e+01	2.881e+00	-7.439	1.79e-11	***
I(Time^2)	5.425e-03	7.252e-04	7.480	1.44e-11	***
factor(Month)1	2.118e+04	2.862e+03	7.401	2.17e-11	***
factor(Month)2	2.118e+04	2.862e+03	7.401	2.17e-11	***
factor(Month)3	2.118e+04	2.862e+03	7.401	2.17e-11	***
factor(Month)4	2.118e+04	2.862e+03	7.401	2.17e-11	***
factor(Month)5	2.118e+04	2.862e+03	7.401	2.17e-11	***
factor(Month)6	2.118e+04	2.862e+03	7.401	2.17e-11	***
factor(Month)7	2.118e+04	2.862e+03	7.401	2.17e-11	***
factor(Month)8	2.118e+04	2.862e+03	7.401	2.17e-11	***
factor(Month)9	2.118e+04	2.862e+03	7.401	2.17e-11	***
factor(Month)10	2.118e+04	2.862e+03	7.401	2.17e-11	***
factor(Month)11	2.118e+04	2.862e+03	7.401	2.17e-11	***
factor(Month)12	2.118e+04	2.862e+03	7.401	2.17e-11	***

2차 추세모형 유효

- 잔차 분석

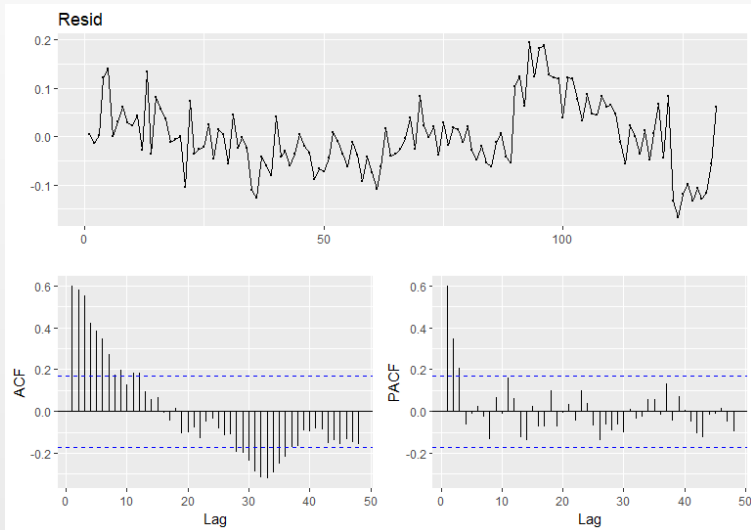
```
> checkresiduals(fit2)
```



- 양의 상관관계 존재

- 오차 모형

```
> Resid <- fit2$residuals  
> ggtsdisplay(Resid, lag.max=48)
```



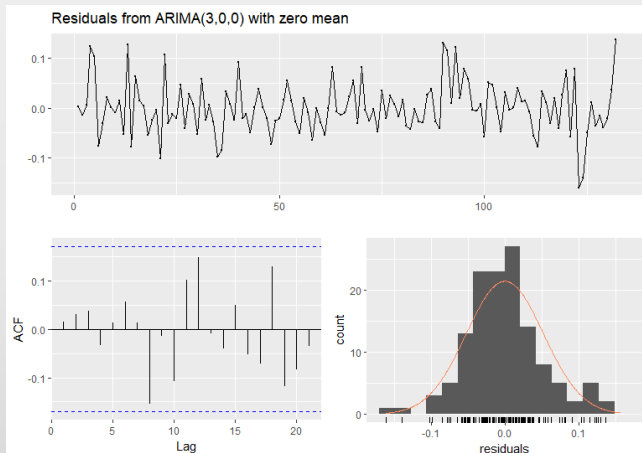
AR(3) 식별

- 오차 모형 적합

```
> fit_r1 <- Arima(Resid,order=c(3,0,0),include.mean=FALSE)
> confint(fit_r1)
                2.5 %      97.5 %
ar1 0.14262044 0.4814755
ar2 0.09449107 0.4352225
ar3 0.03698488 0.3775787
```

```
> checkresiduals(fit_r1)
```

data: Residuals from ARIMA(3,0,0) with zero mean
Q* = 6.0871, df = 7, p-value = 0.5296



```
> confint(Arima(Resid,order=c(3,0,1),
                include.mean=FALSE))
> confint(Arima(Resid,order=c(4,0,0),
                include.mean=FALSE))
```

추가된 모수 비유의적

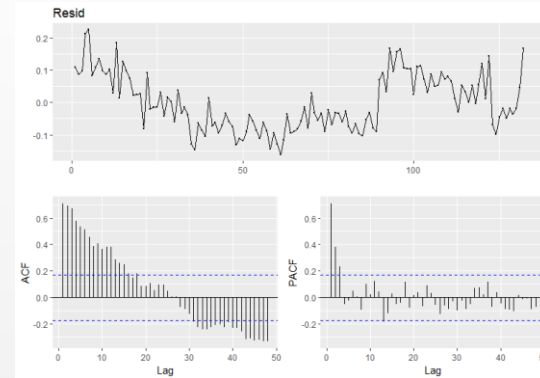
- 2차 추세계절 + AR(3) 오차 회귀모형

```
> fit_x <- model.matrix(fit2)
> f1 <- Arima(tour.ts,order=c(3,0,0),include.mean=FALSE,
              xreg=fit_x,lambda=0)
Error in solve.default(res$hessian * n.used, A) :
  system is computationally singular: reciprocal condition number =
  5.58217e-17
```

- 회귀모형: $Y = X\beta + \varepsilon$
- 모수 추정량: $\hat{\beta} = (X^T X)^{-1} X^T Y$
- 'computationally singular': 역행렬 계산이 안 되는 경우
- 1차 추세 모형을 대신 시도

- 1차 추세 계절 모형의 오차 모형

```
> Resid <- fit1$residuals  
> ggtsdisplay(Resid, lag.max=48)
```



AR(3) 식별

- 오차 모형 추정

```
> fit_r2 <- Arima(Resid, order=c(3,0,0), include.mean=FALSE)  
> confint(fit_r2)
```

```
> checkresiduals(fit_r2)
```

```
> confint(Arima(Resid, order=c(3,0,1),  
               include.mean=FALSE))  
> confint(Arima(Resid, order=c(4,0,0),  
               include.mean=FALSE))
```

오차 모형 AR(3) 확정

- 1차 추세계절 + AR(3) 오차 회귀모형

```
> fit_x <- model.matrix(fit1)
> f1 <- Arima(tour.ts,order=c(3,0,0),include.mean=FALSE,
              xreg=fit_x,lambda=0)
```

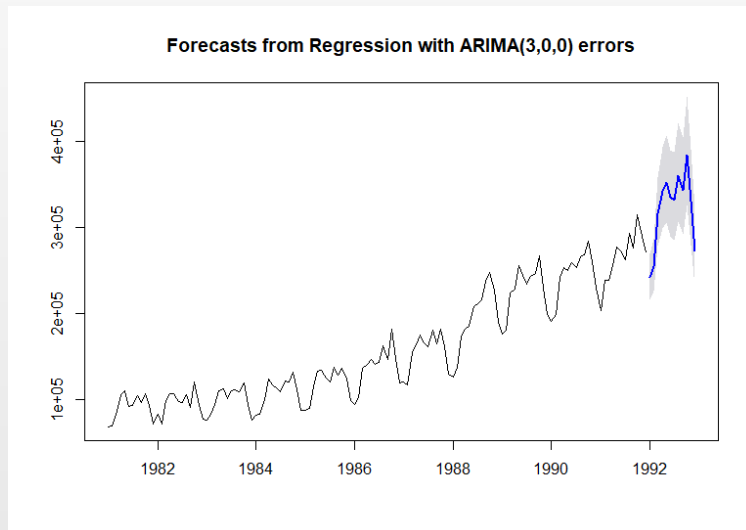
```
> summary(f1)
Series: tour.ts
Regression with ARIMA(3,0,0) errors
Box Cox transformation: lambda= 0

Coefficients:
            ar1      ar2      ar3      Time  factor(Month)1  factor(Month)2
            0.3466  0.2946  0.2382  0.1194            -225.5527            -225.5287
s.e.         0.0873  0.0886  0.0889  0.0092             18.1917             18.1919
            factor(Month)3  factor(Month)4  factor(Month)5  factor(Month)6
            -225.3269            -225.2397            -225.2205            -225.2767
s.e.             18.1920             18.1920             18.1922             18.1923
            factor(Month)7  factor(Month)8  factor(Month)9  factor(Month)10
            -225.2919            -225.2196            -225.2721            -225.1684
s.e.             18.1923             18.1924             18.1925             18.1926
            factor(Month)11  factor(Month)12
            -225.3262            -225.5249
s.e.             18.1927             18.1928

sigma^2 estimated as 0.003213:  log likelihood=199.46
AIC=-364.91   AICc=-359.55   BIC=-315.91
```

- 예측

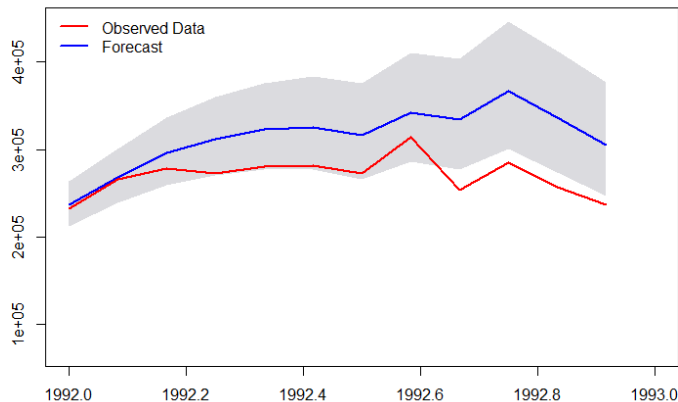
```
> new.t <- time(ts(start=c(1992,1),end=c(1992,12),freq=12))  
> new.x <- cbind(new.t,diag(rep(1,12)))  
> fore_reg <- forecast(f1,xreg=new.x,level=95)  
> plot(fore_reg)
```



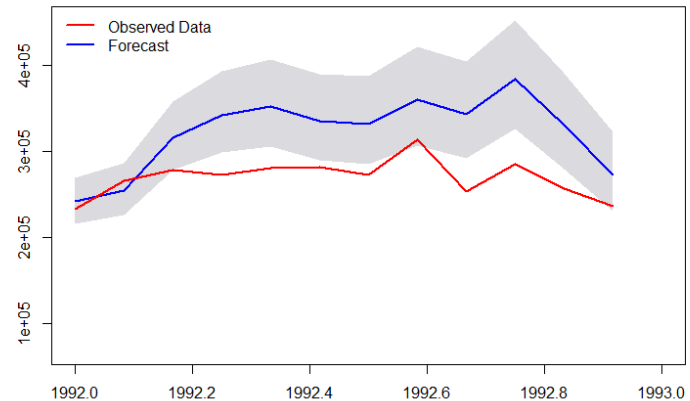
3. 예측결과의 비교

```
> fit4_1 <- Arima(tour.ts,order=c(2,1,0),  
                  seasonal=list(order=c(1,1,0),period=12),lambda=0)  
> fore_arima <- forecast(fit4_1,h=12,level=95)  
  
> f1 <- Arima(tour.ts,order=c(3,0,0),include.mean=FALSE,  
              xreg=fit_x,lambda=0)  
> fore_reg <- forecast(f1,xreg=new.x,level=95)
```

Forecasts from ARIMA(2,1,0)(1,1,0)[12]



Forecasts from Regression with ARIMA(3,0,0) errors



- 예측 정확성 측도 비교

```
> accuracy(fore_arima,tour92)
              ME      RMSE      MAE      MPE      MAPE      MASE
Training set -101.7388  9382.833  6691.828 -0.1666852  4.06514  0.3697756
Test set     -44218.3189 51964.159 44218.319 -16.5746164 16.57462  2.4434060
              ACF1 Theil's U
Training set  0.1129040      NA
Test set      0.6654927  1.920409

> accuracy(fore_reg,tour92)
              ME      RMSE      MAE      MPE      MAPE      MASE
Training set   120.1524  9569.667  6442.998 -0.3146017  4.019269  0.3560258
Test set      -53052.9409 61306.604 54979.697 -19.5490298 20.273884  3.0380559
              ACF1 Theil's U
Training set  0.1121612      NA
Test set      0.5282178  2.260151
```